

U.S. Foreign Policy: Strategic Choice

Branislav L. Slantchev

Department of Political Science, University of California, San Diego

Last updated: January 5, 2016

In order to organize our thinking about foreign policy, we must decide what it is that we want to study and what assumptions we want to make to simplify reality sufficiently to make it comprehensible. As I noted above, our explanations of foreign policy must necessarily boil down to arguments about why particular actors took certain actions, and these arguments must, on average, hinge upon the assumption that somehow each actor was trying to achieve some desired goals. Since no actor, not even the President of the United States, is powerful enough to simply impose its preferred outcomes on others, the defining characteristic of international relations (and so, foreign policy) is the **interaction** among various actors, and it is this interaction that we shall study. At the most abstract level, we must distinguish three components: (i) the actors, (ii) the environment in which they act, and (iii) how outcomes are produced from the actions.

1 The Actors: Preferences and Beliefs

Here are some examples of different actors in whose interaction we might be interested: states fighting a major war, United Nations engaged in peacekeeping operations, governments of two states negotiating a trade treaty, the ministries of a country seeking accession into the European Union, State Department and Department of Defense struggling for control over foreign policy, General Motors and Ford lobbying the government for protection against “unfair” foreign-trade practices, French farmers dumping grapes to protest agricultural policies of the EU, individuals engaging in terrorism.

It should be evident that we are not interested in fixing some particular level of social aggregation as the unit of analysis. That is, we do not want to say that we shall investigate relations between states only, or between leaders of states, or even between organizations within states. International relations are far less conveniently structured than this, and we shall have to account of various different types of actors getting involved.

To deal with this complexity, we shall use an abstract definition of an actor. An actor has two attributes: **preferences** and **beliefs**.

To say that an actor has preferences simply means that it can rank order different outcomes according to some criterion or criteria. For example, consider the situation with Iraq and suppose there are six possible outcomes: (i) Iraq provides acceptable proof of dismantling of its WMD programs, (ii) Iraq agrees to dismantling whatever is left of these programs under international supervision, (iii) Saddam steps down as Iraq’s leader, (iv) the United States invades Iraq and wins, (v) the United States invades Iraq and loses, or (vi) the

US does nothing.

The United States is an actor that has a specific preference ordering. That is, it ranks these alternative outcomes in some rational way. Similarly, we can designate the State Department, or Saddam, or President Bush for that matter as actors, and they all will have their own preference orderings.

The other attribute of an actor is the beliefs it has about the preferences of other actors. Again, since we are interested in interaction among actors, we want to know how these actors will behave, which in turn depends on what they think others will do. To form an expectation about the behavior of other actors, it is necessary to have some belief about what preferences the other actors have. For example, we might be uncertain about whether Saddam's preferences are such that he prefers (i) to (ii) above, but we can hold a belief about the likelihood that it is the case. When actors are uncertain, as it is usually the case because they seldom possess complete information, beliefs are crucial to the choice of action.

Thus, we shall study the interaction among actors, where actors are defined by two attributes, their preferences and their beliefs.

1.1 Unitary and Composite Actors

It is important to understand that actors that we can profitably treat as single "individuals" at a high level of abstraction can themselves be composed of other actors at a somewhat lower level of abstraction. For example, in some contexts, it might be appropriate to define the United States as the actor and postulate some preferences over the risky alternatives. This could be a useful shortcut, and historians often employ it, in some situations: for instance, it might not be too distorting of reality to treat the United States as an actor whose preferences opposed the expansion of Soviet communism during the Cold War. In other contexts, however, this shortcut might be distorting: for instance, we might wish to analyze how the United States would respond to some particular aggressive move by the Soviet Union. Reasonable people can agree on the goal — preventing the success of this move — but disagree about the appropriate course of action. This disagreement can arise because of different political priorities, beliefs about "how the world works", or even organizational and bureaucratic issues.

Here we would need to "disaggregate" the United States into a composite of several relevant actors. But how do we know what these actors have to be? To answer this question, we need to know a bit more how the U.S. foreign policy decision-making process is organized. We shall study this in some detail very soon, so for now let us assume that the important individuals would be the President, the National Security Advisor, the Secretary of State, the Secretary of Defense, and the Chairman of the Joint Chiefs of Staff. We can take these individuals as representing the preferences of the respective organizations they head, which means that they might have very different ideas how the Soviet action might need to be handled. For instance, the JCS chairman might prefer to respond militarily with an action that has the highest chances of success; e.g., a ground invasion. The Defense Secretary might agree with the assessment of the likelihood of success but might be more sensitive to the costs the various actions entail. He might prefer to opt for a much less expensive strategy — e.g., massive air strikes — that might have a smaller probability of success. The State Secretary might be worried about the fallout of using a military option without attempting

a diplomatic solution first. He might prefer to delay the military response until allies could be consulted and the opponent given an opportunity to retreat without an overt confrontation. The National Security Advisor might believe that the Soviet move, while seemingly aggressive, is not actually all that threatening and that even if successful it would not really damage American interests. He might be opposed to any military response but also to any diplomatic intervention which might disturb the allies unnecessarily and give the opponent an opportunity to score points by defying the negotiation attempts. He might think that the appropriate course of action is to do nothing at all and simply ignore the Soviet move. The President might agree with absence of a real threat, but worry about the impact of inaction on the American public; he might believe that the public would never forgive him for failing to resist aggression. Thus, he might want to do something, and that something would have to be more than “merely talking” about a diplomatic solution but definitely less than immediate military action; he might, for instance, decide on a forceful non-military response like a naval blockade.¹

In order to explain the foreign policy choice of the United States government in this scenario, the theory cannot treat the United States as a **unitary actor**. Instead, it will have to incorporate knowledge about the decision-making process at the highest level of government to model the United States as a **composite actor** whose preferences and beliefs are somehow determined by the preferences and beliefs of the five unitary actors we identified. At this point, the theory will confront two issues. The first is merely a repeat of the unitary actor problem we just encountered: even though the Secretary of Defense is an individual, it does not follow that he has to be modeled as a unitary actor; after all, he is the head of a vast, and fairly complex, bureaucratic organization that, at a minimum, comprises structures designed to deal with the three main branches of the military: the Army, the Navy, and the Air Force. When the President asks for advice, the Secretary would focus on the level of forces and manpower necessary to pursue various goals. As the head of this agency, he might be concerned about the appropriate balance among the various branches, their degrees of readiness, and cost effectiveness. He might wish to pursue organizational goals that involve promoting some particular technology at expense of others: e.g., a new stealth bomber instead of more tanks. This organizational goal might bias him in favor of air strikes (that would demonstrate the capabilities of the bomber, and so earn him even more support from the Air Force) and against ground invasion (that would expose the shortage of tanks he had created, and so open him to criticism from the Army). To fully understand the preferences of this actor, one might have to disaggregate it in turn, taking into account the bureaucratic organization of the Defense Department and the individual preferences of the various Secretaries and Under-Secretaries in it. Analogous issue might prompt us to disaggregate the Secretary of State, and so on. Clearly, this type of analysis can become extremely involved and so detailed that it would be nearly impossible to follow. For practical purposes, disaggregation stops at the highest level of abstraction that allows us to make meaningful predictions about the behavior of the composite actor. As before, purpose determines scale and simplification.

¹Students familiar with the 1962 Cuban Missile Crisis will, of course, recognize that this hypothetical example is not fabricated out of thin air.

1.2 Preference Aggregation in Composite Actors

The second issue is the problem of **preference aggregation**. Even though we identified the five actors — which for now we shall treat as unitary — that are of special relevance for the formulation of foreign policy, we have not specified how their preferences and beliefs are aggregated into preferences and beliefs of the composite actor the United States. It could be, of course, that the President acts like a dictator and just implements the action according to his own preferences and beliefs. As we shall see, however, even if the President is ultimately responsible for the final decision, that decision will invariably be shaped by the opinions of those around him. This influence can be informal: the other actors seek to obtain agreement with their preferred action through a process of deliberation and persuasion. The influence can also be formal: the President takes the action that garners the majority vote. Different Presidents will employ different styles of decision-making, and it can run the gamut from near dictators who ignore advice to first-among-equals who carry out the wishes of the majority. They will also surround themselves with different types of individuals, some preferring the company of those whose preferences are not too dissimilar from theirs, and others valuing diversity of opinion.

Suppose that, after intense deliberations all five agree that neither land invasion nor doing nothing are desirable options. They still disagree, however, about the relative merits of air strikes, blockade, and diplomacy. Let’s suppose, for the sake of example, that their individual rankings are as follows:

President	Advisor	State	JCS	Defense
blockade	blockade	diplomacy	air strikes	diplomacy
air strikes	air strikes	blockade	diplomacy	air strikes
diplomacy	diplomacy	air strikes	blockade	blockade

Table 1: Preference orderings of five unitary actors for the composite United States.

Since they cannot persuade each other beyond this, the President decides to use pairwise majority voting. He first asks everyone to choose between blockade and diplomacy. Since three of the five actors prefer diplomacy to blockade, diplomacy is the winner. The President then asks everyone to choose between diplomacy and air strikes. Since three actors prefer the air strikes, the air strikes is the ultimate winner. It appears that the United States prefers air strikes most, followed by diplomacy, followed by blockade. The Chief of the JCS will be happy, but the Secretary of State is distinctly unhappy with this.

Suppose the State Secretary managed to persuade the President to redo the voting but start with the choice between air strikes and blockade. Since three actors prefer blockade to air strikes, the majority winner is blockade, which is then paired with diplomacy. But since three actors prefer diplomacy to blockade, the ultimate winner is diplomacy, which the State Secretary likes a lot. It now appears that the United States prefers diplomacy most, followed by blockade, followed by air strikes. It should already be troubling to you that a “mere technicality” of switching the order of voting has altered the preferences of the composite actor.

It gets worse. The President, who is now saddled with his least preferred option, has

warmed up to the idea of agenda manipulation and decides to redo the voting. He asks everyone to vote on air strikes and diplomacy first. Since three actors prefer air strikes to diplomacy, the winner is air strikes, which is then paired with blockade. Since three actors prefer blockade to air strikes, the ultimate winner is blockade, just what the President wanted. It now appears that the United States prefers blockade most, followed by air strikes, followed by diplomacy.

Thus, depending on the order in which alternatives are considered, using majority voting to determine the preferences of the composite actor from the logically consistent individual preferences of the constituent unitary actors gives us logically inconsistent results, known as **preference cycles**. The United States prefers blockade to air strikes, air strikes to diplomacy, and diplomacy to blockade. These preferences are logically inconsistent because logic dictates that if one prefers blockade to air strikes and air strikes to diplomacy, then one should prefer blockade to diplomacy as well (preferences should be transitive).

The problem with preference cycles is that they make theories unfalsifiable because *every* choice is consistent with the preferences of the composite actor. But if every choice is “rationalized” by these preferences, then we cannot understand why any particular choice was made. It seems that any theory that seeks to rationalize behavior based on preferences is doomed from the start.

1.3 The Need to Consider Institutions

Or maybe not. In fact, our simple example above already suggests one way in which the preferences of the composite actor can be guaranteed to be consistent. If the President acts as the **agenda-setter** and decides the order in which options are brought up for a vote, then he can ensure that the preferences of the United States are exactly the same as his own even though they were ostensibly created by majority voting. Thus, the agenda-setter can not only avoid cycles in aggregate preferences, but can usually ensure that the voting outcomes are very close to his own preferences. This gives agenda-setters considerable power, of course, which is why these formal positions are so desirable when the institutions allow for them. In our case, the President’s elevated rank might informally designate him as the agenda-setter even when there is no formal voting rule in the group of decision-makers he is consulting with. This ability might, in fact, allow us to treat the United States as a unitary actor after all, except in this case its preferences would be those of the President. If, on the other hand, we were interested in the decision-making of another type of government, say a military junta composed of several generals who make collective decisions using majority voting, then we might be able to restrict attention to the general with agenda-setting powers.

The American government system of checks and balances, however, ensures that when it comes to foreign policy, the President might find himself at loggerheads with Congress. The ultimate action the government takes will be based on preferences created by aggregating the preferences of the executive and legislative branches. Congress itself is a very complex institution whose members have to deal with a great variety of possibilities, making the possibility of preference cycles quite distinct. Congress, however, has many rules and practices that eliminate that possibility altogether. Among these institutional features are: (i) the rules of order, which might limit the opportunities for defeated proposals to come back; (ii) reversion points (e.g., preset spending allocations in a budget), which au-

tomatically select an alternative if no proposal receives enough votes to decisively defeat it;² (iii) adoptions of winning alternatives as reversion points, which makes it exceedingly unlikely that voting would cycle back to the original; (iv) committee systems that limit the number of alternatives considered, amendment rules that require that any changes be germane to the committee proposal, or rules that limit the amendments themselves to those proposed by the committee;³ (v) vote-trading practices, which allow a member to exchange a vote on some issue of interest to others for the others' votes on an issue of interest to the member (this allows for the formation of stable winning coalitions); (vi) parties, which restrict the domain of admissible preferences by enforcing party discipline on the members.

The institutional constraints and practices might appear arbitrary and might have somewhat undesirable consequences (e.g., logrolling can produce vastly inflated budgets, and party discipline might polarize Congress resulting in policy deadlock), but they are necessary evils because they impose structure that can induce stability in context where decisions are made by majority rule. This is why we shall often have to consider the institutions in which policy-makers operate, not merely their (imputed) preferences. This is also why we will need to study the process of foreign policy formation in the United States more closely.

1.4 The National Interest

The problem of preference aggregation is much more pressing than our abstract examples might suggest. Consider, for example, the ubiquitous notion of **national interest**, in whose name political leaders and groups purport to act. There are two things here that we should be careful about:

1. How is the national interest determined, and
2. How is the most appropriate action chosen given that national interest?

That is, as a society we probably need to agree on what our common interests are, and once we agree on that, what the best ways to achieve these interests would be.

You have all read history books and are aware of stuff you see on TV. Not a single day goes by without some pundit pontificating on air or in print about the current crises in Iraq and with North Korea, not to mention the perennial Arab-Israeli conflict in the Middle East, the economic difficulties of Latin America, the AIDS epidemic devastating Africa, or the corruption scandals rocking Europe.

All of these discussions are invariably framed in terms of preferences of the participating actors. Historians, journalists, economists, and political scientists are all intensely interested in these preferences because we all look for explanations of behavior by assuming some consistent pursuit of self-interest by these actors. Whether in trying to divine Saddam Hussein's preferences or those of the United States, we all resort to an appeal of instrumentally rational behavior to explain what goes on. ("Instrumentally rational" refers to the assumption that people pursue actions consistent with their goals. That is, people will not willingly hurt their own interests.)

²This also works in law, where the current law stands unless the court explicitly overturns it; the principle of *stare decisis*.

³Under the Closed Rule in the U.S. House of Representatives, no amendments may be offered other than those recommended by the committee itself, which further restricts the range of admissible preferences.

For simplicity, many analysts take the state as the unit of analysis when it comes to important international events. So we talk about a Second Persian Gulf War between America and Iraq, or a crisis between the U.S. and North Korea, or bargaining for more money between Turkey and the U.S. In other words, we often take the state to be the important actor whose behavior we want to explain. It is in this context that you frequently hear the much abused and maligned term “the national interest.” But what is it?

There are several possible ways we can approach the problem, and all of them have been used in international relations theory:

- Objective interest, which overrides all other concerns whether states realize that or not. For example, realism postulates that state survival is the most important national interest and all other goals are subordinated to this one. Liberals tend to argue that the world is not such a dreadful place and that economic well-being is the most important national interest.
- Expression of elite choice. In this view, elites have specific interests that they pursue through the state apparatus, to which they have better access than ordinary people. Elites then “sell” these policies to the rest of us, inducing our choices to conform to their preferences. This works both for democracies and non-democracies (authoritarian or totalitarian regimes).
- Expression of people’s choice. Proponents of democracy argue that the national interest is simply an aggregation of individual preferences. That is, each and every one of us has his or her own preferences. In a democracy, we would then use some aggregation mechanism, usually voting, to arrive at the social preference.

Of course, there is no such entity as a state when it comes to preferences. States do not have preferences, people do. The “objective” interest is really a simplifying assumption in the tradition that treats states as actors in their own right. It is also fairly narrow because it only specifies what it takes to be the most important objective — security or power or wealth — and therefore may not provide much of a guidance when we want to deal with less apocalyptic issues. Still, there are many venerable schools of thought — which you will encounter in this course — that insist that we need not look below the abstract level of the state, or, if we do, we need not go very deep at all. Structural realism is among the former while classical realism, Marxism, and liberalism are among the latter.

The other two ways of looking at the national interest may be more helpful. Instead of postulating an objective to an abstract entity (the state), we take the national interest to be really an expression of individual preferences, whether they are elite decision-making groups or voters. In these views, a state implements the “best” policy consistent with either elite or voter preferences. The approaches tend to disagree as to who gets to decide what’s “best” and whose preferences the policies will tend to reflect: those of the majority voters, of the few powerful members of the elite? However, they agree that somehow some relevant group of people has to agree on what the national interest is and how to get at it.

People have disagreements, usually vehement, on both of these issues. For example, you and I may disagree whether maintaining stable international markets is ultimately in our national interest. I, being internationally minded, may strongly believe that of America

fails to keep the economy stable, it will eventually cause enormous problems domestically as well. You, being a firm agnostic about the value of globalization, may maintain that this is nonsense, and America should rely on its huge internal market and perhaps insulate itself as much as possible. There are many contentious issues in foreign policy, and what constitutes the national interest is a question that is seldom answered, although many talking heads seem to assume that it is self-evident. Quickly: is preventing the spread of Islamic fundamentalism in the national interest? Or plugging the ozone hole? Or saving hundreds of thousands AIDS victims in Africa? Or assisting Israel against the Palestinians? Or the Palestinians against Israel? Or championing women's rights in Afghanistan? Which is more important? What about making sure Pakistan doesn't sell nuclear technology to other unsavory characters besides Lybia's ex-strongman Muammar Qaddafi? Or that Russia keeps its precocious bio warfare specialists from selling their services for hard currency? Or preventing Russia from gobbling up parts of neighboring countries? The list is potentially endless.

Suppose, however, that somehow we, as society, agree on what constitutes the national interest. For example, we all agree that America should strive to keep the global economy stable. We then fall into the next pit: what is the best way to do this? Should we maintain close links with repugnant regimes like the Saudi Arabia's autocrats just because they sit on the world's largest oil reserves that our European and Asian friends need so badly? Should we pursue a more hard-line policy in the Middle East to secure our ability to react to potential problems when the unpopular regimes eventually fall apart, as they must inevitably do? Or maybe we should hike up gas prices domestically so people don't drive needlessly? Or maybe we should invest heavily in fossil fuel-efficient technologies or even totally new hydrogen-based ones? Or perhaps tax the hell out of gas-guzzling SUVs that no sane person should be driving anyway? Or maybe everyone who thinks that Americans should be limited in their ability to drive tanks on highways is a goddamn pink Commie bastard that we should get rid of? New Yorkers and Bostonians with their nice public transportation and city lives that involve walking from place to place may be inclined to support policies that make driving costlier. But Californians and Texans who are rather spread out and who commute long distances may be much less enthusiastic. Anyway, even if we agree on the ultimate goal, we may still disagree ferociously on the methods we should use to get there.

Given all these disagreements that are bound to result from the simple fact that people are different, hold disparate beliefs, perceive the world in various ways, and have differential access to the levers of government, we should either appoint a dictator who simply implements the choices she wants (and hopefully these would be the ones she believes are for the good of the many) or else we must find a way to aggregate our disparate opinions into some sort of **collective choice**. (Note that even if we are ruled by a small elite, an oligarchy of sorts, then the members of this ruling elite must still find a way to aggregate *their* preferences into choices that the smaller collective body will make.) We have, however, already encountered a fundamental problem with group decision-making in the abstract setting above. Whether it is the elites or the voters who get to define the national interest or the means of achieving it, each group has to arrive at some ranking of alternatives and pick the one it likes best, and we now know that the institutional features of the group can be crucial in determining what group preferences will look like.

In other words, *it is impossible to conceive of the national interest solely in terms of preferences of the individuals that comprise the polity, even if these individuals are restricted to privileged elites. The national interest will depend on the institutional characteristics of the government, which themselves usually evolve after years of contentious politics, and thus tend to reflect the distribution of power in society.*

2 The Environment: Actions and Informational Structure

Actors do not make their choices in vacuum. The other defining component of our approach to international relations is the strategic environment in which interaction takes place. An environment is composed of **actions** that are available to the actors and an **information structure**.

The first is simply the set of actions which summarize how actors can interact. For example, during crisis negotiations, the set of actions might include (i) escalating the crisis by taking a provocative step, such as mobilizing troops or sending aircraft carriers into a volatile region, (ii) deescalating a crisis, (iii) starting a war, (iv) backing down and accepting the other side's demands, (v) producing new demands, (vi) insisting on previous demand and adopting a wait-and-see attitude, (vii) organize support of allies, (viii) make an offer on an unrelated issue linked to the opponent accepting your position on the one currently under consideration. The list can go on and on, although in most cases it is surprisingly short because it excludes all "irrelevant" choices. For example, although an actor may choose to produce more sugar, this choice will not be part of the crisis bargaining environment because it is not relevant for the decisions to be made in that strategic context. The environment limits the possible actions physically as well. For example, the action "initiate nuclear strike" is simply not available to non-nuclear powers.

The second component of the environment is its information structure. That is, what the actors can know and what they have to infer from observable behavior of others. This is related to beliefs because that information available in the environment determines in part the beliefs that the actors will hold. For example, suppose that in the crisis one side ostensibly deploys an armored division in an attempt to force the other to accept its demands. The move may appear aggressive, causing the other to update its beliefs and revise its estimate of the likelihood that its opponent is prepared to go to war. However, suppose that from its spies that side also learns that the tanks are old and there is insufficient fuel and supplies to actually put them in action. The deployment now appears as an empty bluff, and so the revised beliefs will very likely be different.

Thus, the actors (preferences and beliefs) interact in strategic environments (actions and information).

3 Strategic Interaction

Now, notice that I said "strategic" environment. What do I mean by **strategic interaction**? While we have defined the actors and the environment they operate in, we have not specified how outcomes are produced from their actions. The crucial aspect of interaction is that outcomes are not the result of any one actor's choices. Instead, in international relations, the choices of many actors determine outcomes.

An actor cannot choose an action simply because it has the best direct effect on the outcome it wants. Rather, it has to take into account the choices of others because they also affect the final outcome. So, an actor will choose an action both for the action's direct effect and its indirect effect on the actions of others. International politics is all about interdependent decision-making. That is, each actors does his best to further its goals knowing that the other actors are doing the same.

To give you a flavor of some of the issues involved, consider the social problem called the **Prisoner's Dilemma**. This problem involves two actors who must decide whether they want to cooperate with each other or not. This game has four possible outcomes, they both cooperate, $\langle C, C \rangle$, only player 1 does, profileCD, only player 2 does, profileDC, and neither does, $\langle D, D \rangle$. Assume that each player's most preferred outcome is when only the other player cooperates, the second most preferred outcome is when both cooperate, the next to last outcome is when both defect, and the least preferred outcome is when he cooperates but the other player does not. For example, suppose the actors are states and "not cooperate" refers to implementing a protectionist economic policy (e.g., imposing a tariff on all goods imported from the other actor), whereas "cooperate" refers to maintaining free trade policies. Then, each player likes it best when it runs a protectionist policy itself (income from the tariffs and protecting competing domestic producers) but the opponent maintains an open regime (so the player's exports are sold on the opponent's country). Free trade is the next best regime, followed by a "tariff war" in which both countries impose tariffs that stifle trade. The worst outcome is to maintain an open regime while the opponent engages in protectionism.

	Protectionism	>	Free Trade	>	Tariff War	>	Open Policy
Country 1	(D, C)	>	(C, C)	>	(D, D)	>	(C, D)
Country 2	(C, D)	>	(C, C)	>	(D, D)	>	(D, C)
<i>Payoffs</i>	4	>	3	>	2	>	0

Table 2: Preferences in the Prisoner's Dilemma.

If you look at the preference orderings, you will see that each player's most preferred outcome is the other player's least preferred one. You might reasonably conclude that neither of these outcomes would be sustainable because the player who is supposed to cooperate unilaterally would instead impose a tariff as well. Since free trade is the second-best outcome for both players, you might then conclude that this should be the outcome produced by rational play. Unfortunately, this will not be the case: if a player believes that his opponent will choose to cooperate, then he is strictly better off not cooperating. In fact, not cooperating is the dominant strategy in this scenario: it is always the best option for each player regardless of what the other player does. This means that the only rationalizable outcome is $\langle D, D \rangle$, the tariff war.

Pause for a minute to think what this means. We have a social situation in which both players agree that cooperating with each other is the second-best choice for both of them. Unfortunately, pursuing their individually rational strategies makes both players worse off. Rationality (at least in this sense) condemns the actors to their next-to-last preferred outcome. In this instance, they will engage in a costly tariff war that will make both of them worse off relative to the free trade regime. They did not do this because they were stupid, ir-

rational, or mistaken. They did this because their incentives in this situation are not aligned properly to support mutual cooperation.

This simple example (with somewhat startling implications) is an instance of “strategic interaction,” which can easily become quite complex because it involves forming expectations about what other actors are going to do, which in turn depends on what they think you are going to do, which, of course, depends on what you think they think you are going to do, and so on and so forth. Going through the chain of reasoning can be pretty difficult because you will end up in an infinite “I think that you think that I think that you think . . .” regression.

In most general terms, *we are going to explain behavior by rationalizing it; and we rationalize it by showing how each actor pursues the best possible strategy given his preferences over outcomes and his expectations about what the other actors are doing. We shall assume that each actor does this, and as a result each will choose the best possible strategy and expect everyone else to do the same. When all actors in a strategic interaction choose their best possible strategies, nobody will be willing to deviate to an alternative strategy, and thus everyone’s expectations will also be rationalized. Because nobody wishes to changes to a worse strategy, the set comprising one optimal strategy for each player is called an **equilibrium** (more specifically, a Nash equilibrium, after the man who invented it). Our definition of **rational behavior** is that it corresponds to an equilibrium strategy. In other words, if we can construct a reasonable model and find that an actor’s behavior we observe is consistent with an equilibrium strategy in this model, we can explain that behavior by saying that it is the best way to pursue the actor’s interests given what every other involved actor is doing. In this way we can make sense of that behavior and the (often unpleasant) outcomes it produces.*

The tool for analysis of strategic interaction is called **game theory**, and it developed as a branch of applied mathematics early in the 20th century, but went nowhere until the US government financed researched for national security purposes in the mid 1960s. It was from these studies initiated for the purpose of finding ways of dealing with the Soviet Union that researchers discovered methods of dealing with uncertainty, beliefs, and strategic interaction in a productive way. In 1994, the Nobel prize in Economics went to three game theorists, the mathematician John Nash, the economist Reinhard Selten, and the strategic theorist John Harsanyi. The models that we are going to use in this class are all game-theoretic (although I will not teach you game theory itself).